

EXAMEN 1EVA – UD2 FICHEROS

XML → CSV y JSON

Ejercicio 2 – (4 ptos)

En el periódico "**El Mundo**" desean almacenar en archivos CSV y JSON los contenidos de RSS de noticias producidos. Desean que el equipo de programadores procese los archivos RSS de la Web y almacenen algunos datos de las noticias.

Implementar el proyecto "**Ejercicio2**" que construya el fichero "**noticias.csv**" y "**noticias.json**" a partir de los datos almacenados en el archivo "**deportes.xml**" en formato RSS.

(el profesor entregará el archivo completo).

El archivo XML se debe leer con la librería **DOM**

El archivo JSON se generará con la librería **javacsv** y utilizando la clase **CsvWriter**

El archivo JSON se generará con la librería **gson**

De la información que contiene el archivo XML en formato RSS sólo almacenaremos el **titular** (title), el **link** y las **categorías**.

El archivo rss contiene los datos de un canal (channel) que en nuestro caso es el de Deportes. Este canal contiene muchas noticias dentro de las marcas "<item>". La aplicación deberá extraer los campos solicitados y almacenarlos en objetos de Java.

En Java se deberán crear Clases que definan los objetos necesarios para el proceso.

Entrada: *deportes.xml*

```

<rss version="2.0">
  <channel>
    <title>Deportes // elmundo</title>
    <link>http://www.elmundo.es</link>
    <description>Deportes // elmundo</description>
    <language>es</language>
    <copyright>(c) 2019, Unidad Editorial, S.A.</copyright>
    <pubDate>Sat, 26 Oct 2019 11:25:29 +0200</pubDate>
    <lastBuildDate>Sat, 26 Oct 2019 11:25:29 +0200</lastBuildDate>
    <category>News and Politics</category>
    <ttl>60</ttl>
    <atom:link href="http://www.elmundo.es/rss/deportes.xml"
      rel="self" type="application/rss+xml"/>
    <image>
      <title>Deportes // elmundo</title>
      <url>http://estaticos.elmundo.es/imagen/canalima144.gif</url>
      <link>http://www.elmundo.es</link>
      <width>144</width>
      <height>24</height>
      <description>elmundo.es</description>
    </image>

    <item>
      <title>¿Por qué pita el Wanda a Simeone y a su capitán?</title>
      <description>
        El Atlético recibe este sábado al Athletic (21.00 h.) en su estadio,
        incómodo por el debate en torno a las figuras del capitán y
        el entrenador.
        <a href="https://www.elmundo.es/deportes/futbol/primera-division/2019/10/26/5db30fddfdddfdfa8e8b4612.html">Leer</a>
        
      </description>

      <dc:creator>CARLOS GUISASOLA</dc:creator>
      <link>https://www.elmundo.es/deportes/futbol/primera-division/2019/10/26/5db30fddfdddfdfa8e8b4612.html</link>
      <category>Koke Resurrección</category>
      <category>Raúl García</category>
      <category>Joao Félix</category>
      <category>Gabi</category>
      <category>Atlético de Madrid</category>
      <category>Copa del Rey</category>
      <category>deportes</category>
      <category>Primera División</category>
      <category>fútbol</category>

```

```

<category>Athletic de Bilbao</category>
<media:description type="html">
    El Atlético recibe este sábado al Athletic (21.00 h.) en su estadio, incómodo por el debate en torno
    a las figuras del capitán y el entrenador.
</media:description>
<media:title type="html">LALIGA</media:title>
<media:content url="https://e00-elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/26/15720784152247.jpg"
    medium="image" width="487" height="325"/>
<media:thumbnail url="https://e00-elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/26/15720784152247_150x0.jpg"
    width="150" height="100"/>
<guid>https://www.elmundo.es/deportes/futbol/primera-division/2019/10/26/5db30fddfdfffa8e8b4612.html</guid>
<pubDate>Sat, 26 Oct 2019 08:43:52 +0200</pubDate>
</item>

<item>
<title>El pequeño reino de Piqué en Andorra</title>
<description>
    <a href="https://www.elmundo.es/deportes/futbol/2019/10/24/5db1d93521efa0f1678b4584.html">Leer</a>
    
</description>
<dc:creator>FRANCISCO CABEZAS</dc:creator>
<link>https://www.elmundo.es/deportes/futbol/2019/10/24/5db1d93521efa0f1678b4584.html</link>
<category>Gerard Piqué</category>
<category>deportes</category>
<category>FC Barcelona</category>
<category>fútbol</category>
<media:description type="html">
    Los gritos y las risas retumban contra la puerta del vestuario. Junto a los bancos y las taquillas
    hay colocada una mesa, sobre la que reposa un tapete verde. A su alrededor van ti
</media:description>
<media:title type="html">FÚTBOL</media:title>
<media:content url="https://e00-elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/25/15720173657142.jpg"
    medium="image" width="487" height="325"/>
<media:thumbnail url="https://e00-elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/25/15720173657142_150x0.jpg"
    width="150" height="100"/>
<guid>https://www.elmundo.es/deportes/futbol/2019/10/24/5db1d93521efa0f1678b4584.html</guid>
<pubDate>Thu, 24 Oct 2019 22:38:53 +0200</pubDate>
</item>

*** Resto de registros ***
</channel>
</rss>

```

Una vez cargado el archivo en objetos de Java, se realiza un proceso para guardar esta información en un archivo CSV con el siguiente formato:

Salida: noticias.csv

```
titular|link|categorias|imagen
Titular de la noticia|https://dominio/path/recurso.html|[ca|categoria1, categoria2, categoria3, ... ]|https://dominio/path/recurso.jpg
Titular de la noticia|https://dominio/path/recurso.html|[ca|categoria1, categoria2, categoria3, ... ]|https://dominio/path/recurso.jpg
Titular de la noticia|https://dominio/path/recurso.html|[ca|categoria1, categoria2, categoria3, ... ]|https://dominio/path/recurso.jpg
Titular de la noticia|https://dominio/path/recurso.html|[ca|categoria1, categoria2, categoria3, ... ]|https://dominio/path/recurso.jpg
```

En nuestro ejemplo, la salida esperada es:

Salida: noticias.csv

```
titular|link|categorias|imagen
¿Por qué pita el Wanda a Simeone y a su
capitán?|https://www.elmundo.es/deportes/futbol/primera-division/2019/10/26/5db30fddfdfffa8e8b4612.html|[Koke Resurrección, Raúl
García, Joao Félix, Gabi, Atlético de Madrid, Copa del Rey, deportes, Primera División, fútbol, Athletic de Bilbao]|https://e00-
elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/26/15720784152247.jpg
El pequeño reino de Piqué en Andorra|https://www.elmundo.es/deportes/futbol/2019/10/24/5db1d93521efa0f1678b4584.html|[Gerard Piqué,
deportes, FC Barcelona, fútbol]|https://e00-elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/25/15720173657142.jpg
...
```

Además, aprovechando que tenemos las clases y los objetos con sus datos, crearemos un archivo JSON cuyo resultado esperado es:

Salida: *noticias.json*

```
{
  "noticias": [
    {
      "Titular": "¿Por qué pita el Wanda a Simeone y a su capitán?",
      "Link": "https://www.elmundo.es/deportes/futbol/primera-division/2019/10/26/5db30fddfdddfdfa8e8b4612.html",
      "categorias": [
        "Koke Resurrección",
        "Raúl García",
        "Joao Félix",
        "Gabi",
        "Atlético de Madrid",
        "Copa del Rey",
        "deportes",
        "Primera División",
        "fútbol",
        "Athletic de Bilbao"
      ],
      "imagen": "https://e00-elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/26/15720784152247.jpg"
    },
    {
      "Titular": "El pequeño reino de Piqué en Andorra",
      "Link": "https://www.elmundo.es/deportes/futbol/2019/10/24/5db1d93521efa0f1678b4584.html",
      "categorias": [
        "Gerard Piqué",
        "deportes",
        "FC Barcelona",
        "fútbol"
      ],
      "imagen": "https://e00-elmundo.uecdn.es/assets/multimedia/imagenes/2019/10/25/15720173657142.jpg"
    },
    *** Resto de registros ***
  ]
}
```

CRITERIOS DE CALIFICACIÓN

De forma general se valorará positivamente:

- Que la estructura del proyecto tenga las carpetas **lib** para las librerías y **datos** para los ficheros
- Que las clases estén almacenadas en un Java Package denominado **clases**
- Que las clases sean serializables y definan todos sus métodos: constructores, getters, setters y toString
- Documentación con comentarios en el fuente

Versiones	Contenido del Proyecto	Puntos
V.1	Lectura del XML, + recorrer mostrando los elementos solicitados con println	1 Ptos
V.2	V.1 + recorrer cargando los elementos en objetos	2 ptos
V.3	V.2 + Crear fichero CSV	3 ptos
V.4	V.3 + Crear fichero Json	4 ptos